Feature-supported Multi-hypothesis Framework for Multi-object Tracking using Kalman Filter

Saira Saleem Pathan, Ayoub Al-Hamadi, Mahmoud Elmezain, Bernd Michaelis

Institute for Electronics, Signal Processing and Communications (IESK),
Otto-von-Guericke-University Magdeburg
Germany,
{Saira.Pathan|Ayoub.Al-Hamadi}@ovgu.de

ABSTRACT

A Kalman filter is a recursive estimator and has widely been used for tracking objects. However, unsatisfying tracking of moving objects is observed under complex situations (i.e. inter-object merge and split) which are challenging for classical Kalman filter. This paper describes a multi-hypothesis framework based on multiple features for tracking the moving objects under complex situations using Kalman Tracker. In this framework, a hypothesis (i.e. merge, split, new) is generated on the basis of contextual association probability which identifies the status of the moving objects in the respective occurrences. The association among the moving objects is computed by multi-featured similarity criteria which include spatial size, color and trajectory. Color similarity probability is computed by the correlation-weighted histogram intersection (CWHI). The similarity probabilities of the size and the trajectory are computed and combined with the fused color correlation. The accumulated association probability results in online hypothesis generation. This hypothesis assists Kalman tracker when complex situations appear in real-time tracking (i.e. traffic surveillance, pedestrian tracking). Our algorithm achieves robust tracking with 97.3% accuracy, and 0.07% covariance error in different real-time scenarios.

Keywords

Multi-object tracking, Traffic surveillance, Applications, Image Processing

1. INTRODUCTION

In computer vision research, tracking algorithms have great significance. This is due to the existing complexities in object tracking such as object appearance, illumination variation, and shadow. Besides, inter-object merge and split are the main issues when tracking multiple moving objects. These issues undermine the performance and efficiency of the tracking algorithms.

Tracking using Kalman filter has been extensively studied during last decades. A wide range of literature is available but Maybecks [May79] provides a comprehensive exposure to Kalman filtering in his paper; whereas a detail review of Kalman filter in visual tracking is given in [Fun03]. Nguyen et al. [Ngu03] used Kalman filter in distributed tracking system for tracking moving people in a room which are monitored by different cameras. Chang et al. [Cha01] used both Bayesian network and Kalman filtering to solve the problem of correspondence between multiple objects. In [Yu04], a video surveillance system is proposed where detection, recognition and tracking of the objects are done. More recently Czyzewski and Dalka [Czy08] used Kalman filter with traditional RGB color-based approach to measure the similarity between moving objects. The above mentioned methods used various data association techniques to handle the complex situations with Kalman tracker. However, single feature-assisted criterion may fail in background clutter, inter-object merge and split. Multiple-hypothesis frameworks supported by multiple features are more robust because it assures generation of correct contextual information of moving objects under diverse situations.

For feature-assisted tracking a good object descriptor is essential. A well known color-based matching technique is the RGB color histogram [Swa90]. This approach has been used with various tracking algorithms [Col05][Yan05]. Besides, color-based tracking methods are also combined with statistical and stochastic approaches. Limin [Lim04] used color-based stochastic algorithm for object tracking, whereas [Pér02] tracks objects in cluttered environments using hue-saturation histogram with the particle filter based probabilistic technique.

However, these techniques are not robust to illumination variation. In addition to color-based techniques, object's features such as, points and lines are exploited to track vehicles [Bey97, Sta99]. Also in [Pet99] active contour models of the vehicles are extracted for tracking purpose. However, feature-based tracking is sensitive to application and therefore an accurate initialization is required to extract the reliable features.

Recently, many hybrid approaches have been proposed where multiple features are combined to improve the robustness over the single feature-based methods, for example [Isa98]. However, these methods [Kum07] [Xio04] have shown the advantage of using multiple feature-supported tracking using various trackers such as the mean-shift estimator and the particle filter respectively. However, the selection and the integration of object's features are still fuzzy.

In this paper we have proposed a multi-hypothesis framework where the online hypothesis is generated on the basis of the contextual association probability. We proposed a multi-featured (i.e. spatial size, color and trajectory) similarity approach in which the color similarity is obtained by the proposed correlation-weighted histogram intersection (CWHI). The statistical similarity of the size and the trajectory is computed using general method. These size and trajectory similarity quantities are multiplied with the fused normalized color correlation. A hypothesis is generated on the basis of the accumulated association probability which identifies the status of moving object when inter-object merge or split is observed and assist Kalman tracker to continue tracking.

This hypothesis is generated online and needs no prior training, proving the efficiency of the proposed framework. An overview of general tracking system is shown in Fig.1. This illustration describes the complex situation and the generation of respective hypothesis during the real-time tracking.

This paper is organized as follows. Section 2 provides a detailed description of the proposed tracking algorithms. Experimental results are presented and discussed in section 3. Section 4 sums up with the concluding remarks including the future research direction.

2. ALGORITHM DESCRIPTION

Kalman filter is an optimal estimator that predicts and corrects the states of the linear processes, such as vehicle tracking or spacecraft. It is not only practically smart but attractive in theoretical terms as well. However, an accurate model of the process is an essential requirement for efficient performance.

Following is the mathematical description of Kalman Filter [Wel95]. We consider a tracking system, where

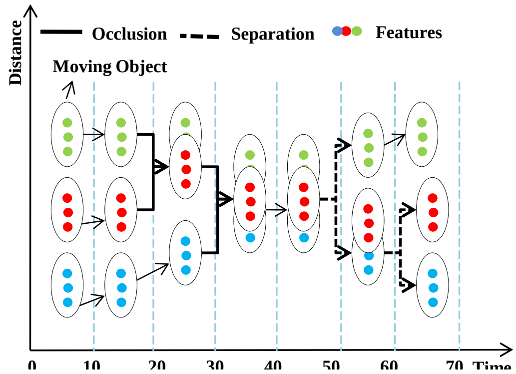

Figure 1. Simplified overview of tracking concept using Kalman Filter which describes the confusing situation (inter-object split and merge).

x_k is the state vector that represents the dynamic behavior of the target with the subscript k indicating discrete time. The objective is to estimate x_k from the set of the observed data z_k. The Kalman filter is described by the following set of equations.

- **Process Equation**

$$x_k = \mathbf{A}\, x_{k-1} + w_{k-1} \quad (1)$$

$$p(w) \sim N(0, Q) \quad (2)$$

Where **A** represents the transition matrix and x_k is the state at time *k-1* to time *k*; where w_{k-1} is the Gaussian process noise $N(.)$ with normal probability distribution $p(w)$.

- **Measurement Equation**

$$z_k = \mathbf{H}\, x_k + v_k \quad (3)$$

$$p(v) \sim N(0, R) \quad (4)$$

Where **H** is the measurement matrix and z_k is the measurement observed at time *k-1* to time *k* and v_k is the Gaussian measurement noise N (.) with normal probability distribution $p(v)$.

- **Time Update Equations**

Equation (1) and (3) describes a linear model at time *k*. As x_k is not measured directly therefore the information provided by the measurement z_k is used to update the unknown state x_k. Apriori estimate of state $\hat{x}_k^-$ and covariance error estimate P_k^- is obtained for the next time step *k*.

$$\hat{x}_k^- = \mathbf{A}\hat{x}_{k-1} + w_k \quad (5)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (6)$$

- **Measurement Update Equations**

These equations are associated with the feedback of the system. The objective is to estimate aposteriori estimation $\hat{x}_k$ which is a linear combination of the apriori estimate $\hat{x}_k^-$ and the new measurement z_k. The equations are given below:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (7)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (8)$$

$$P_k = (I - K_k H) P_k^- \quad (9)$$

Where K_k is the Kalman gain, $\hat{x}_k$ is the estimated state of the target object with covariance error P_k.

The time and measurement equation's pair are recursively repeated with the previous aposterior estimate to predict the new aprior estimate. This recursive behavior of estimating states is one of the highlights of the Kalman filter.

2.1 Object Tracking using Kalman Filter

For tracking objects using Kalman filter, a motion model is required. This model is defined in terms of its states, measurement updates and noise. The system consists of multi-tracker where each tracker is associated with a moving object which enters in the scene as illustrated in Fig 1. The state vector consists of center of gravity of the object trajectories t_k^x and t_k^y at time *k* and *k-1*.

$$x_k = \begin{bmatrix} t_k^x \\ t_k^y \end{bmatrix} \quad (10)$$

In order to assure tracking, we consider two states which represent the trajectory of moving objects. However, Accumulated Contextual Association Probability (*ACAP*) which is described in section 2.3 is used to guide Kalman tracker in complex situations. The measurement vector of the system adopts the following from:

$$z_k = \left[t_k^x \; t_k^y \right]^T \quad (11)$$

Where **A** is the transition, and **H** is the measurement matrix of our system with associated noise $w_k$and v_k:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Finally, we can write the equations for our tracking system.

$$x_k = A x_{k-1} + w_k \quad (12)$$

$$z_k = H\, x_k + v_k \quad (13)$$

Tracking the moving object has several phases. Each newly detected object is assigned a new tracker with associated states (x_k and z_k). In the next frame, normal updating state is done until any complex situation arises. When the inter-object split or merge appears our tracker system follows the generated online hypothesis which is based on *ACAP* to help the tracker to stick with its own object.

The main motivation is to track multiple objects efficiently using Kalman filter with the proposed *ACAP* based approach.

2.2 Multi-Hypothesis Framework

In this paper, a multi-hypothesis framework is proposed which is based on multiple-feature similarity criteria.

2.2.1 Color Similarity

We have proposed a CWHI (correlation-weighted histogram intersection) technique where the idea is adapted from Histogram Intersection (HI) which is proposed in [Swa90]; and with a recently proposed technique by Jia et al. [Jia06] for car number plate matching. In this approach, Gaussian weights are calculated to develop a relationship among the distance differences of the histograms for the moving object. However, for tracking multiple moving objects, this technique is not computationally efficient.

In our approach, a fused color correlation is used as weights and applied to conventional HI. This describes the relationship between the color distance and the normalized correlation weights. A color histogram is constructed using the hue and the saturation values of the moving object. Result shows that the incorporation of the correlation weights is efficient to find the association of moving objects under confusions.

The color histogram of each moving object is extracted by calculating hue and saturation of each pixel from the RGB values and then binning them to create a histogram. The hue and saturation values of the histogram are calculated for every moving object and to compute the similarity between them.

Mathematically, we first compute the normalized correlation of hue channel ρ_{hue} and normalized correlation saturation ρ_{sat} separately. Where Cov_{Th} and Cov_{Mh} are the covariance of the hue values of the object at time *k* and *k-1* respectively. Similar, convention is used for Cov_{Ts} and Cov_{Ms}. Also, σ_{ThMh} and σ_{TsMs} are the standard deviations for the hue and the saturation of both the objects at time *k* and *k-1*. These values are then combined to generate the fused color correlation ρ_{fused} which is used with HI to compute the $CWHI$.

$$\rho_{hue} = \frac{\sigma_{ThMh}}{Cov_{Th} Cov_{Mh}} \quad (14)$$

$$\rho_{sat} = \frac{\sigma_{TsMs}}{Cov_{Ts} Cov_{Ms}} \quad (15)$$

$$\rho_{fused} = \rho_{hue} \times \rho_{sat} \quad (16)$$

$$CWHI = \sum_{i=his_M} \sum_{j=his_T} \min\left(h_M(i), h_T(j)\right) exp^{\left(-\frac{d}{2\rho_{fused}^2}\right)} \quad (17)$$

The color distance d in (17) is calculated using the following equation:

$$d = \sqrt{(h_M - h_T)^2 + (s_M - s_T)^2} \quad (18)$$

 ISBN 978-80-86943-93-0

Where h_M represents the hue of the object at time *k-1* and h_T represents the hue of the object at time *k*. Similar convention is used for saturation that is s_M and s_T respectively.

2.2.2. *Size Similarity*

The size similarity probability is calculated at every frame in video sequence which is then combined with the fused color correlation (ρ_{fused}). The product of these two shows the contextual association with respect to the size of the moving object. It is calculated as follows:

$$S_{size} = \rho_{fused} \times P_{size} \tag{19}$$

Where P_{size} represent the similarity probability and is calculated by taking the size distribution of all moving objects with each moving object.

2.2.3. *Trajectory Similarity*

The trajectory similarity probability of the moving object is calculated at every frame of the video sequence along with the corresponding normalized correlation. The product S_{traj} shows the contextual association with respect of moving object's trajectory. Following is the formulation of the trajectory similarity:

$$S_{traj} = \rho_{fused} \times P_{traj} \tag{20}$$

Where, ρ_{fused} and P_{traj} represents the fused color correlation and the similarity probability of the trajectory, respectively. The trajectory's similarity probability is calculated by finding the trajectory distribution of moving objects with the individual object.

2.3 ACAP for Hypothesis Generation

The ranks for respective hypothesis is calculated and generated after the accumulation of multi-feature similarity probabilities. On the basis of ranks, the hypothesis is generated for respective moving object on the fly. The contextual association probability is calculated using the following equation:

$$ACAP = CWHI \times S_{size} \times S_{traj} \tag{21}$$

Where, *ACAP* represents the current contextual association of the moving objects with respect to its previous occurrences when inter-object merge occur.

In Figure 4, the graphs show the results of tri-feature based contextual association probabilities of moving objects. It is shown that the moving object identities can easily be managed using multi-feature based criterion.

3. EXPERIMENTAL RESULTS

The experimental result shows the performance of the proposed framework. The proposed approach is tested on both the real and the synthetic videos. The sample recording is taken from IESK video analysis repository[1].

In Fig. 2, synthetic video is shown in which three virtual cars are moving in different direction. When two cars are merged (i.e. inter-object occlusion), the trajectory of the occluding moving object is connected (merged) with the occluder. It is observed that in frame *k+15*, the two cars (i.e. labeled as blue and red) are merged with the occluder object (i.e. middle car). Whereas, when two merged moving objects split, the overlapped track is also separated into two tracks. The generation of hypothesis (i.e. split or merge) is based on the result of *ACAP* where the weak *ACAP* shows the merged moving object (i.e. combined with other object) and strong *ACAP* shows the occluder object (i.e. hide or impede other moving objects).

Implementation on the real video sequences is presented in Fig.3; it is observed that tracking of moving objects is possible after inter-object merge and split. Identities of both the moving objects are managed through *ACAP* hypothesis generation criteria. However, errors are noticed during the initial states of merging and splitting. It is also observed that both merge and split occurs at irregular interval of time.

Fig 4(a) shows the Kalman tracking results. The graphs show the result (i.e. synthetic video) of our *ACAP* in Fig. 4(b) where the moving object association represents the strong *ACAP* during tracking and managed the respective hypothesis online. The presented work is based on our initial research analysis on multi-object tracking using stochastic tracking algorithm and features supported multi-hypothesis techniques.

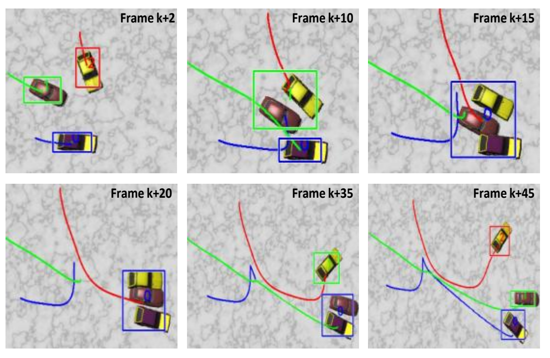

Figure 2. Tracking results of the moving cars with motion trajectories are presented. Motion trajectories are showing that the tracking of objects using Kalman tracker is possible despite of inter-object merge and split due motion.

[1] Institute of Electronics, Signal Processing and Communications, OvG University, Magdeburg, Germany.

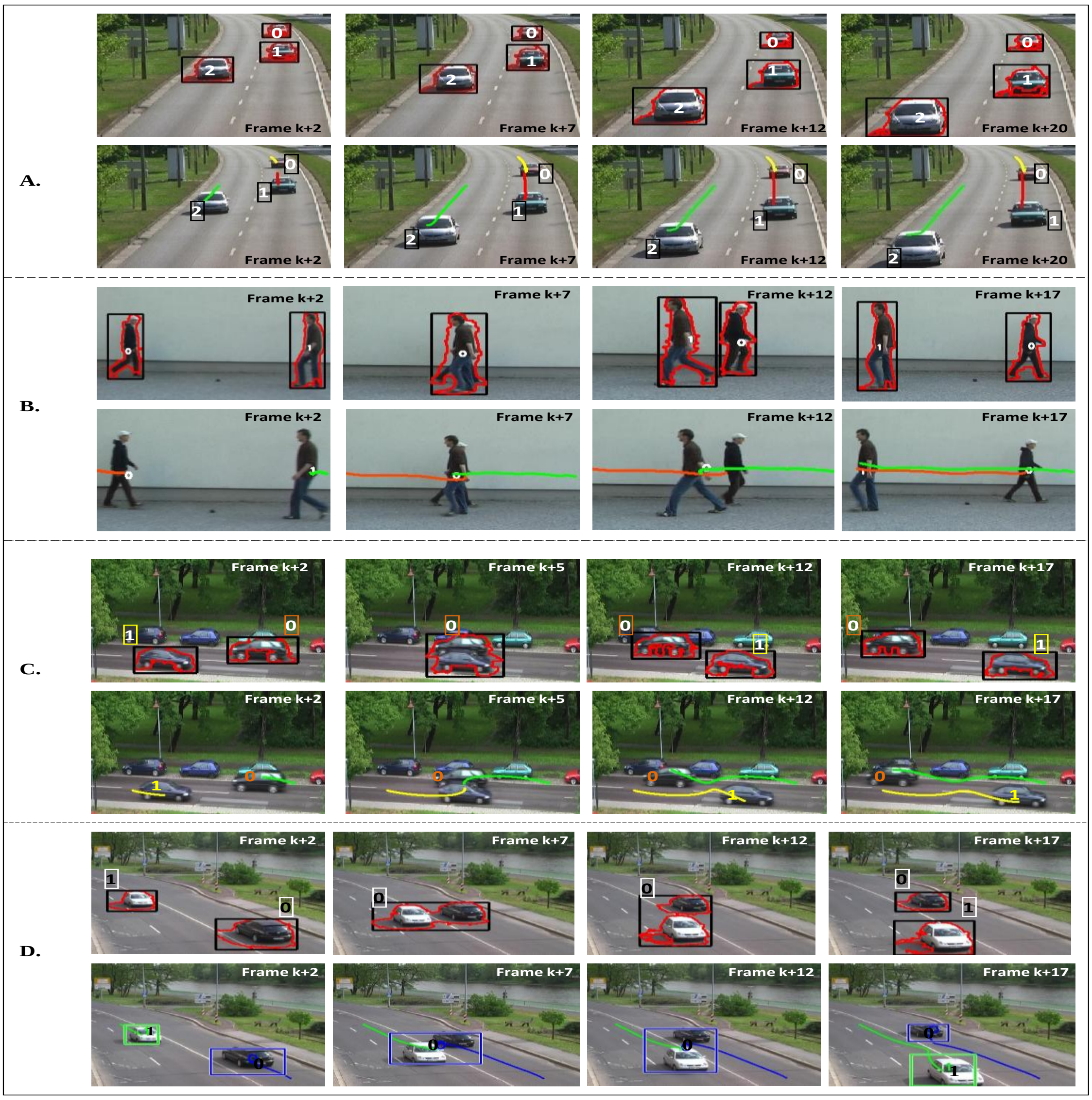

Figure 3. Analysis results of moving cars with motion trajectories (trajectories are associated with Index Numbers). In A, the car are tracked through the indexes and shows the tracking in normal case. In B, two moving people are fully occluded in the frame k+7. It is observed that the objects are tracked during merge and split. In C, two moving cars with same color are traced after confusion because we are not relying on single feature. So, we are able to maintain the identities of objects and assist Tracker. In D, the merge occurs due to shadows but the *ACAP* assisted Kalman tracker is able to track both the cars during the complex situations.

These results motivate us to further investigate the feature-based matching techniques together with stochastic tacking algorithms. Particularly, more distinctive techniques such as, fuzzy logic with learning classifier will be investigated. The result of proposed algorithm shows good tracking with 97.3% accuracy and 0.07% covariance error.

4. CONCLUSION

A multi-hypothesis assisted framework is proposed for tracking multi-object in complex situations (i.e. inter-object split and merge). A multi-featured corresponding hypothesis (i.e. split, merge, and new) on the basis of accumulated contextual association probability. The graphs show that the individual moving objects can be identified via maximum similarity probability during inter-object merge and split. Future research will be focused on investigating fuzzy approaches with learning algorithm such as AdaBoost or Support Vector Machine (SVM) techniques in order to test large surveillance videos with complicated real situations. The proposed approach has shown good tracking results with contextual association of the object identities on test videos containing confusing situations.

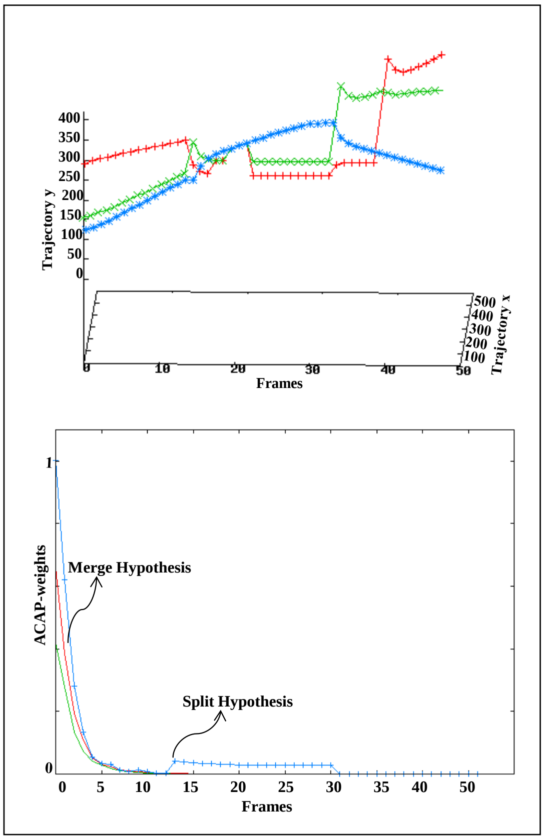

Figure 4. (a, b) Graphs show the result of multi-hypothesis based tracking using Kalman Filter. The first plot represents the tracking trajectories by the Kalman tracker, whereas, the second graph represents the *ACAP* data association hypothesis during the complex situations. It shows clearly that the identity of moving object is managed during the complex situations.

5. ACKNOWLEDGMENTS

This work was supported by DFG (Tracer MI 377/12-2); and DFG grants (SFB/TR 62, TP C1).

REFERENCES

[Bey97] Beymer, D., Mclauchlan, P., Coifman, B., Malik, J.: A real-time computer vision system for measuring traffic parameters. Conference of Computer Vision and Pattern Recognition, pp. 495–501, 1997.

[Cha01] Chang, T. H., and Gong, S.: Tracking Multiple People with a Multi-Camera System. IEEE Workshop on Multi-Object Tracking, pp. 19–26. 2001,

[Col05] Collins, R., Liu, Y., Leordeanu, M.: Online selection of discriminative tracking features. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, pp. 1631–1643, 2005.

[Czy08] Czyzewski, A., Dalka, P.: Examining Kalman Filters Applied to Tracking Objects in Motion. 9th International Workshop on Image Analysis for Multimedia Interactive Services, pp.175-178, 2008.

[Fun03] Funk, N.: A Study of the Kalman Filter applied to Visual Tracking. University of Alberta, Project for CMPUT 652, 2003.

[Isa98] Isard, M., Blake, A.: Icondensation: Unifying low-level and high level tracking in a stochastic framework. European Conference on Computer Vision,Vol. 1, pp. 893–908,1998.

[Jia06] Jia, W., Zhang, H., He, X., Wu. Q.: Symmetric colour ratio gradients in spiral architecture.18th International Conference on Pattern Recognition, 2006.

[Kum07] Kumar, P. S., Guha, P., Mukerjee, A.: Colour and Feature Based Multiple Object Tracking Under Heavy Occlusions. International Conference on Advances in Pattern Recognition, 2007.

[Lim04] Limin,X.: Object tracking using color-based Kalman particle filters. IEEE International Conference on Signal Processing, pp. 679–682, 2004.

[May79] Maybeck, P. S.: Stochastic Models Estimation and Control. New York, Academic Press, Vol. 1, 1979.

[Ngu03] Nguyen, N., Bui, H. H., Venkatesh, S., West, G.: Multiple camera coordination in a surveillance system. ACTA Automatica Sinica, Vol. 29, pp. 408-422, 2003.

[Pér02] Pérez,P., Hue, C., Vermaak, J., Gangnet, M.: Color-Based Probabilistic Tracking. 7th European Conf. on Computer Vision, pp. 661–674, 2002.

[Pet99] Peterfreund, N.: Robust tracking of position and velocity with Kalman snakes. IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 21, pp.564–569, 1999.]

[Sta99] Stauffer, C., Grimson, W. E. L.: Adaptive background mixture models for real-time tracking. Conf., Computer Vision and Pattern Recognition, Vol. 2, pp. 246–252, June 1999.

[Swa90] Swain, .M, Ballard, D.: Color Indexing. Third IEEE International Conference on Computer Vision, Japan, pp. 11–32, 1990.

[Wel95] Welch,G., Bishop, G.: An Introduction to the Kalman Filter. University of North Carolina at Chapel Hill, Technical Report 95-041, 1995.

[Xio04] Xiong,T., Debrunner, C.H.: Stochastic car tracking with line- and color-based features. Journal of ITS, Vol 5, pages 324-328, 2004.

[Yan05]Yang,T., Li, S. Z., Pan, Q., Li, J.: Real-time multiple objects tracking with occlusion handling in dynamic scenes. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 970– 975, 2005.

[Yu04] YU, H., Wang, Y., Kuang, F., Wan, Q.: Multi-moving Targets Detecting and Tracking in a Surveillance System. 5th World Congress on Intelligent Control and Automation, China, June, 2004.